




ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Latent structure of executive function in preschoolers: A systematic review and meta-analysis

Alessia Bandettini ^{*} , David Giofrè, Sabrina Panesi, Sergio Morra, Laura Traverso

Department of Education Sciences, University of Genoa, Italy

ARTICLE INFO

Keywords:

Systematic review
Meta-analysis
Executive function
Preschool
Latent factor

ABSTRACT

In the last decades, an impressive amount of research demonstrated the importance of the early development of executive function for concurrent and subsequent psychological development and adjustment. Nevertheless, the structure of executive function in this age range is still a matter of debate. The present systematic review and meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement and focused on studies that used confirmatory factor analysis with at least two indicators for identifying executive function components in preschool and toddler children. In addition to highlighting methodological issues and points of convergence and divergence across studies through the systematic review (sample size, age range, type of indicators, relationship between indicators and latent variables), the meta-analysis provided a quantitative synthesis of the correlations between the most frequently studied latent constructs. In particular, we found a substantial correlation ($r = .66$) between working memory and inhibition, suggesting that while these functions are closely related in early development, they remain distinguishable components of EF. To move forward in understanding early executive function development, future research should benefit from taking into account some of the issues identified.

The term Executive Function (EF) is often described as an "umbrella" term encompassing a set of high-order cognitive abilities crucial for self-regulation and organizing complex, goal-directed behaviors. These abilities include initiating, planning, organizing an action, shifting attention between different tasks, and suppressing prepotent responses to reach functional goals. These domain-general processes have an adaptive function, playing a crucial role when we are faced with new situations or problems where automatic or impulsive responses would be inadequate or ineffective (Diamond, 2013).

EF plays a crucial role from the earliest stages of development. Although these cognitive processes continue to mature throughout adolescence, the most rapid developmental gains are typically observed between the ages of 3 and 5 years (Carlson, 2005). During this period, children exhibit a reduction in impulsive responses, and greater use of self-regulatory strategies. Moreover, preschool development of EF has been linked to improvements in a range of socio-emotional competencies, including theory of mind, problem-solving abilities, and school readiness (Blair & Razza, 2007; Hughes & Ensor, 2007; Senn, Espy, & Kaufmann, 2004).

Despite a large number of studies, there is no consensus on the model that best describes the EF structure. Historically, models considered EF as a unitary factor. Baddeley (1986; 2003), for example, proposed a "Central Executive", defined as an attentional system with limited capacity, involved in storing and processing information, focusing and shifting attention, reasoning and decision-making. The Central Executive in turn refers to the Supervisory Attention System (SAS), defined by Norman and Shallice

* Correspondence to: Department of Education Sciences, University of Genoa, Corso Andrea Podesta', 2, Genoa (GE) 16121, Italy.
E-mail address: alessia.bandettini@edu.unige.it (A. Bandettini).

<https://doi.org/10.1016/j.cogdev.2025.101612>

Received 11 December 2024; Received in revised form 31 July 2025; Accepted 3 August 2025

Available online 25 August 2025

0885-2014/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

(1980; 1986) as an attentional control system related to individual differences in cognitive tasks that require planning and control skills. At the same time, dissociations among EF task performance observed in clinical context supported the emergence of multi-component models of EF (Godefroy, Cabaret, Petit-Chenal, Pruvo, & Rousseaux, 1999; Stuss & Benson 1984; Tranel, Anderson, & Benton, 1994). In the debate about the unitary versus multicomponent nature of EF, a fundamental contribution was given by the seminal paper of Miyake and colleagues (Miyake et al., 2000). These authors used confirmatory factor analysis (CFA) to assess the fit of several EF models in an adult sample. They compared a model positing a unitary central executive against other ones portraying the central executive as comprising various executive dimensions. Through CFA, Miyake and colleagues identified three interrelated but distinct EFs: inhibition—the ability to suppress task irrelevant cognitive processing and ignore salient yet irrelevant features of the environment (Miyake et al., 2000); shifting—the ability to switch between different operations or levels of processing (Miyake et al., 2000); and updating—the capacity to encode, hold and monitor incoming information in working memory (WM), replacing outdated information with new information (Morris & Jones, 1990). According to this perspective, EF appears to rely on WM, which involves the simultaneous maintenance and manipulation of information. Furthermore, EF seems to involve a common inhibitory component, which would explain why these functions are interrelated yet distinct.

While multiple EF models continue to exist (e.g., Zelazo, Carter, Reznick, & Frye, 1997), the three-component model identified by Miyake and colleagues has become the most widely accepted theoretical framework for understanding EF (Friedman & Miyake, 2017; Miyake & Friedman, 2012). However, consensus on the best model in adulthood does not necessarily extend to childhood, where findings are less consistent. Lehto, Juujarvi, Kooistra, and Pulkinnen (2003) found an executive structure similar to that proposed by Miyake et al. (2000) in children aged from 8–13 years, identifying three latent dimensions interpreted as inhibition, WM, and shifting. In contrast, Huizinga, Dolan, and van der Molen (2006) and van der Sluis, de Jong, and van der Leij (2007) identified only two EF components, WM and shifting, and did not identify a separate inhibition factor, possibly due to an inappropriate selection of inhibition tasks. A two-factor model was also proposed by Shing, Lindenberger, Diamond, Li, and Davidson (2010), who identified only two dimensions (WM and inhibitory control), in children aged 10–15 years. Lee, Bull, and Ho (2013) also found a two-factor structure including an updating and a combined inhibition-switching factor in children aged 5–13 years.

The scenario among preschool children shows similarly discordant results. Some studies (e.g., Hughes, Ensor, Wilson, & Graham, 2009; Shing et al., 2010; Wiebe, Espy, & Charak, 2008; Wiebe et al. 2011) suggested that a single factor model best described the EF structure in preschoolers. Other research indicated that a two-factor model better represented the EF structure in this age group. Miller, Giesbrecht, Müller, McInerney, and Kerns (2012), for example, identified a two-factor model consisting of WM and inhibition in a sample of children aged 3–5 years. Similarly, Lerner and Lonigan (2014) claimed in favor of a model distinguishing between WM and inhibition in children aged from 3.5 to 5 years. This was confirmed by another study, performed by Usai, Viterbori, Traverso, and De Franchis (2014), also identifying a two-factor model in children between 5 and 6 years. In this model inhibition was a separate factor, while WM and shifting were merged into a single factor (see also Monette, Bigras, & Lafrenière, 2015). Differently, Panesi and Morra (2020) reported a similar two-factor model, including a WM and a combined inhibition – shifting – updating factors.

A systematic review by Karr and colleagues (2018) revealed that most of the studies on preschool children identified one or two factors, whereas studies in older children more frequently identified two or three factors. The inconsistency in the developmental literature regarding the structure of EF can arise from several causes. First, it is important to note the variability in how WM and updating are conceptualized across different age groups. While in adult populations there is a consensus that these are separate constructs (Himi, Bühner, & Hilbert, 2021), each contributing uniquely to the EF framework, in the developmental literature the distinction between WM and updating is often blurred (Morra, Panesi, Traverso & Usai, 2018; Panesi, Bandettini, Traverso & Morra, 2022). For instance, Garon, Bryson, and Smith (2008) in their review of EF in preschoolers, identified inhibition, shifting, and WM as key EFs, actually equating WM with the updating processes described by Miyake et al. (2000). Similarly, Diamond (2013) considered inhibition, WM, and cognitive flexibility as core EFs that develop in early childhood, where WM often encompasses the functions of updating. This conflation can lead to confusion and inconsistencies in the measurement and interpretation of EFs in young children. The assumption that WM subsumes updating tasks not only impacts the theoretical structure of EF models but also affects the empirical findings regarding their development. Understanding these distinctions and their implications is crucial for accurately assessing cognitive development and tailoring educational strategies accordingly.

Another source of inconsistency in the developmental literature on EF pertains to the variability in constructs considered across different studies. For example, Panesi and Morra (2020) included tasks assessing WM, inhibition, shifting, and updating in their models of EF structure. In contrast, many studies focus more narrowly on a subset of constructs, typically assessing only WM and inhibition (e.g., Lerner & Lonigan, 2014; Wiebe et al., 2011). Nevertheless, it should be noted that EF include several higher-order cognitive processes such as planning, and problem-solving (Diamond, 2013) which are less investigated. Additionally, distinctions made by authors like Zelazo and Müller (2002) between ‘cool’ EF and ‘hot’ EF further complicate the issue. These authors categorize cool EF to involve cognitive skills like inhibition and WM, whereas hot EF includes processes linked to emotional and motivational aspects, such as delay of gratification and affective decision-making. The considered EF influence the research design and, consequently, the results.

Inhibition, WM, and shifting are commonly investigated due to their central role in cognitive development. Additionally, they appear relatively circumscribed and, therefore, they seem easy to operationalize (see Miyake et al., 2000). Nevertheless, operationalizing these EF for children is far from simple. EF tasks require diverse knowledge and skills (i.e. impurity problem, Miyake et al., 2000), that children are acquiring (Hughes & Graham, 2002). In the literature, while multiple inhibitory tasks are available, shifting and updating tasks are fewer, which affects study design (Wiebe et al., 2008). Moreover, to use CFA, EF tasks should be very different and very basic in non-EF requests and very similar in the EF required. Nevertheless, some studies use tasks that seem characterized by a different complexity level. The selection of tasks to be used as indicators of each EF could also impact or bias the findings. For instance, if the tasks chosen as indicators of inhibition place also a high load on the young participants’ WM capacity, this could inflate the

correlation between the WM and inhibition latent variables or even render them undistinguishable.

Other problems depend on the use of different numbers of indicators to identify specific EF components in different studies. Some authors considered two or more indicators for each function considered (e.g., Usai et al., 2014); others employed a different number of indicators depending on the investigated function (Schoemaker et al., 2012); still others used less than two indicators (Garon, Smith, & Bryson, 2014) or used multiple indicators from a single task (Hughes, Ensor, Wilson, & Graham, 2009). This can be problematic, as using a limited number of indicators per factor might bias the results of latent factor models.

Furthermore, the label "preschool" is used in the literature to encompass different age ranges. For example, in some studies this term is used to refer to children from 3 to 6 (Lonigan, Lerner, Goodrich, Farrington, & Allan, 2016; Miller, Giesbrecht, Muller, McInerney, & Kerns, 2012; Miller, Müller, Giesbrecht, Carpendale, & Kerns, 2013; Völter et al., 2022); in others is used to refer to children from 2–6 years (Wiebe et al., 2008). The inclusion of a broad age range in studies of preschool children's EF can pose significant challenges due to the rapid developmental changes of early childhood. Some reviews have documented substantial differences in EF task performance among children aged 2, 3, 4 and 5 years; these differences are likely reflective of distinct developmental trajectories (e.g., Carlson, 2005; Best & Miller, 2010).

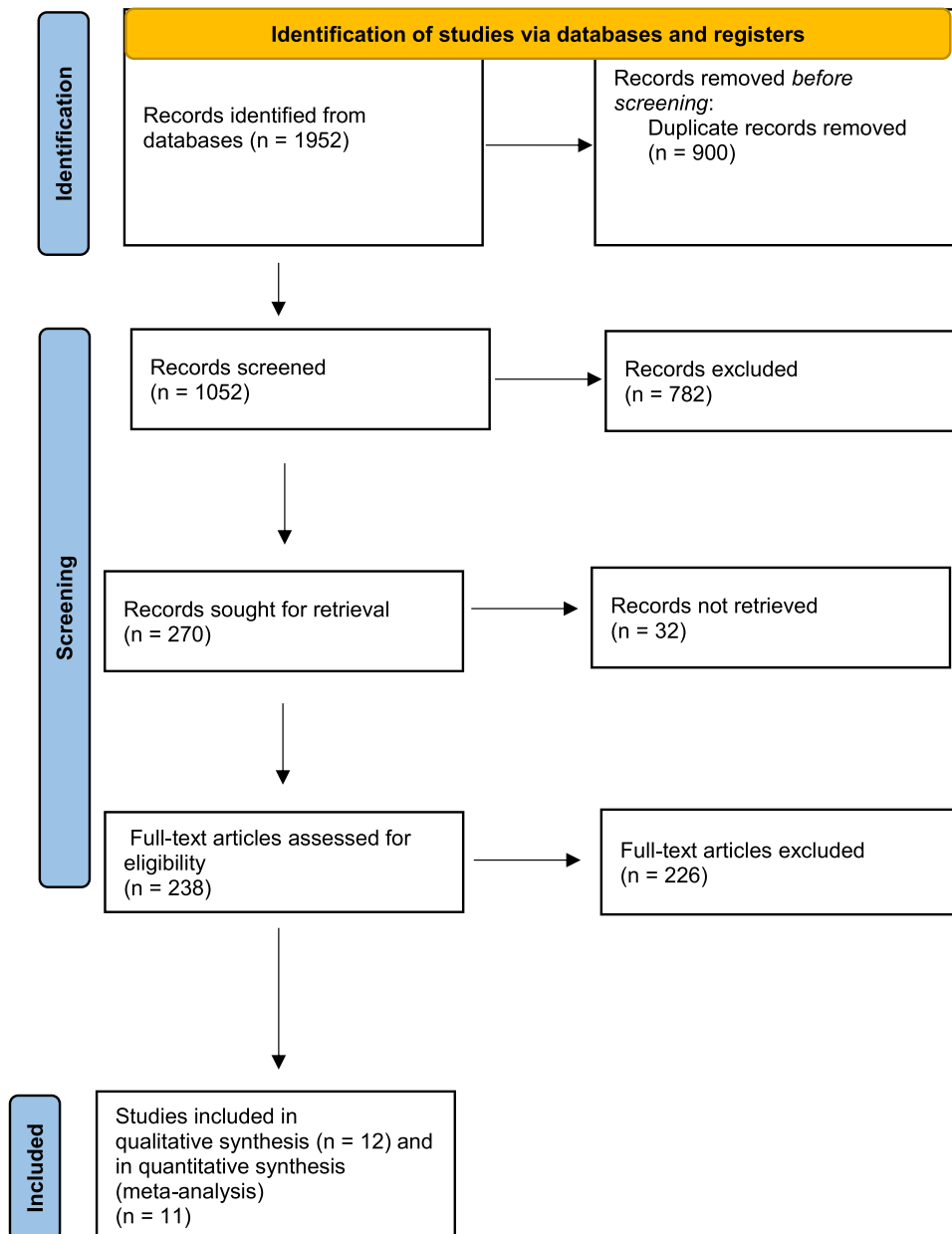


Fig. 1. Search and selection process.

All these inconsistencies and methodological problems are a key factor contributing to the inconsistencies observed in developmental studies of EF, particularly in the preschool years where cognitive abilities are rapidly developing. Although existing evidence points to both quantitative and qualitative changes in EF during early development, the exact nature of these functions remains somewhat elusive.

In light of these challenges, it is crucial to comprehensively review and understand the methodologies employed in these studies. [Karr et al. \(2018\)](#) included preschoolers in their systematic review but did not focus exclusively on this age group. Their research aimed to evaluate the empirical support for various EF models and to identify the best fitting models based on CFA results. Our study builds upon these foundations with a more focused examination of the methodologies, sample characteristics, and indicator features used in preschool EF research. By doing so, with the present systematic review we aim to uncover the reasons behind the observed inconsistencies and provide insights into how future research can be structured to better understand EF development in early childhood.

This review is structured around several key research questions: (a) which theoretical model received more empirical confirmation; (b) which models were compared and contrasted; (c) which latent variables were examined; (d) how the tasks employed were actually informative of the constructs they are intended to measure; (e) how the age range of participants affects the resulting model. Our review aims to not only to compare different EF models, but also to critically analyze the methodologies used to identify them.

Moreover, we also aimed to investigate the relationship between latent variables with a meta analysis. Specifically, our goal was to go beyond the descriptive synthesis offered by the systematic review and provide a more robust, quantitative contribution by analyzing the empirical data reported in the included studies. To conduct these analyses, we focus on the latent variables that, according to the results of the systematic literature review, have been the most frequently studied, analyzing the strength of the correlations between them, in order to assess how these constructs relate to each other in preschool children. In this way, the meta-analysis offers a more empirical examination by statistically analysing the data available from these studies, providing a deeper understanding of the latent structure of EF in preschool children.

To advance our understanding of EF structure during preschool years, examining systematically the methodologies and findings of previous studies is essential. Such efforts will not only clarify the developmental trajectory of EF but also enhance our understanding and help to develop future and more informed studies.

1. Method

Concerning the systematic review, it adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement (Moher, Liberati, Tetzlaff, Altman, & the PRISMA Group, 2009). The flowchart shown in [Fig. 1](#) outlines the search and selection process.

Given that EF encompasses multiple higher-order cognitive processes, various theoretical models have been proposed to conceptualize its structure. In this review, we decided to focus on WM, inhibition, shifting, and updating as described by Miyake and colleagues (2000), as these constructs are prominently used in the literature.

Moreover, we decided to limit our review to studies that utilized performance-based measures of EF, excluding those that relied solely on questionnaire-based assessments. This decision was guided by evidence indicating a weak correlation between performance-based and questionnaire-based evaluations of EF ([Conklin, Salorio, & Slomine, 2008](#); [McAuley, Chen, Goos, Schachar, & Crosbie, 2010](#); [Sølsnes, Skranes, Brubakk, & Løhaugen, 2014](#)).

In addition, our focus was specifically on studies investigating the latent structure of EF through CFA, a method widely recognized as the most effective approach for validating EF structure ([Miyake et al., 2000](#)). As a theory-driven multivariate analysis technique, CFA plays a crucial role in defining the organizational framework of EF.

Finally, given the inconsistencies in the literature regarding the definition of the preschool age range and the potential impact of this variability on findings, we chose to adopt the broadest possible age range that can be considered preschool (2–6 years). This decision minimizes the risk of excluding studies that could provide valuable insights for our research objectives.

1.1. Literature search

The literature search was conducted on 8th November 2022, including the following databases: PsycINFO, PsycArticles, Psychology and Behavioral Sciences Collection, and Pubmed. To minimize the risk of publication bias, grey literature from ProQuest, which also includes dissertations, was also considered. The search was limited to publications in English. Keywords used across all databases included combinations of: "executive function*" or "cognitive flexibility" or "shifting" or "inhibition" or "inhibitory control" or "updating" AND "working memory" or "M capacity" or "attentional capacity" or "memory span" AND "confirmatory factor analysis" or "CFA" or "SEM" or "Structural Equation Model" or "model" or "factor" or "structure" AND "child*" or "preschool*" or "kindergarten". In PsycINFO, PsycArticles, and PubMed, the selected keywords could appear in any field; however, in ProQuest, these keywords were only required to be present in the abstract or summary text.

1.2. Abstract screening

From the initial search a total of 1952 abstracts were identified. After removing 900 duplicates, 1052 papers remained. Two authors independently and blindly reviewed 263 articles, corresponding to 25 % of total abstracts. Agreement between reviewers was 100 %. Titles and abstracts were reviewed to identify the relevant articles. The selection process followed a set of inclusion criteria: (1) articles published on scientific journals or dissertations; (2) empirical studies; (3) participants were humans; (4) participants were children

under 7 years; (5) samples consisted of typically developing children; (6) Inhibition, shifting, updating and WM were assessed through performance-based tasks; (7) more than one EF were assessed. Articles that did not provide sufficient information to determine inclusion were carried forward to the full-text screening phase.

1.3. Full-text articles screening

A total of 270 references proceeded to full-text screening, from which 238 full-text articles were retrieved. For their assessment, we added three inclusion criteria: (1) CFA was used; (2) at least two models were compared; (3) each EF was measured by at least two tasks. Twelve papers met all inclusion criteria, corresponding to 14 studies as two articles included two studies each (Hume, 2015; Miller et al., 2012).

Concerning the meta-analysis, from the set of studies that we included in the systematic review, we focused on the ones which investigated the functions of WM and inhibition calculating latent correlations between these two constructs. Therefore, we needed to exclude Mulder and colleagues' study (2014), since they classified EF as cool and hot and they did not take into account WM and inhibition. Furthermore, Hume's article (2015) included separate analysis for children aged 3 and 4. Finally, we decided to include Panesi and Morra's work (2021) because they employed WM tasks alongside with measures of inhibition, even if they were considered tapping the same underlying EF factor in the original report.

Studies indicating a latent correlation of one or exceeding one were not included in our analysis (Hume, 2015 –children aged 4; Völter et al., 2022; Wiebe et al., 2008), since is not possible to calculate standard errors and variances under these circumstances. Furthermore, every latent correlation in the meta-analysis was calculated based on the correlation matrix presented in the original report. The final dataset consisted of nine latent correlations between the inhibition and WM factors. Following the guidelines provided by Borenstein, Hedges, Higgins, and Rothstein (2009), the correlations were converted to Fisher's z' score, and then converted back to the correlation metric. A random effect model was used. Standard errors were calculated using the formula $1/(N-3)$. The "metafor" package for R (Viechtbauer, 2010) was used for the analyses.

2. Results

Table 1 shows the characteristics of the samples recruited for the selected studies. As observed, the samples considered across the reviewed studies appear relatively homogeneous in terms of gender distribution. In nearly all cases, approximately half of the participants were female. Slightly lower proportions of female participants—below 40 %—were observed only in Miller et al. (2012), Miller et al. (2013), and Panesi and Morra (2021). Moreover, all studies were conducted in Western, Educated, Industrialized, Rich, and Democratic (WEIRD) countries, specifically in North America and Western Europe. Ethnic composition is more difficult to analyse systematically, as several studies did not report this information (e.g., Monette et al., 2015; Mulder, Hoofs, Verhagen, van der Veen, & Leseman, 2014; Panesi and Morra, 2021; Usai et al., 2014), while others provided data only on the majority ethnic group (e.g., Miller et al., 2012; Miller et al., 2013; Völter et al., 2022). Overall, the samples tend to include a majority of White/Caucasian participants. In contrast, greater heterogeneity was observed in terms of sample size and age range, variables whose role in influencing the resulting EF

Table 1
Sample characteristics.

References	Sample, N	Age (range)	% Female	Nationality	Ethnicity
Hume (2015)	117	40.98 (36–45)	49.4	USA	65.4 % White, 19.8 % African American, 4.6 % Hispanic, 2.1 % Asian, 8.0 % other ethnicities
Lerner & Lonigan (2014)	289	55.74 (45–63)	53	USA	31 % African American/Black, 57 % Caucasian/White, 3 % Hispanic/Latino, 3 % Asian, and 7 % other ethnicities
Lonigan et al. (2016)	241	54.23 (38–69)	53.5	USA	51.5 % Latino/White, 0.8 % Black, 0.4 % Asian, 0.4 % Hawaiian/Pacific Islander, 7.1 % multi-racial, 39.8 % not reported
Miller et al. (2012)	129	50.45 (36–68)	39.5	Canada	80 % Caucasian
Miller et al. (2013)	129	50.04 (36–68)	39.5	Canada	80 % Caucasian
Monette et al. (2015)	275	68.43 (/)	54.5	Canada	/
Mulder et al. (2014)	2437	28 (20–37)	49	Netherlands	/
Panesi & Morra (2021)	78	31.1 (25–37)	34.6	Italy	/
Usai et al. (2014)	175	68.5 (63–76)	43.4	Italy	/
Völter et al. (2022)	190	49.65 (36–72)	54.2	UK	Predominantly White
Wiebe et al. (2008)	243	47 (28–72)	55.6	/	171 Caucasian, 43 African American, 9 Asian American, 1 Native American, 4 Hispanic, 14 multi-racial, 1 not reported
Wiebe et al. (2011)	228	37 (/)	50.4	/	173 Caucasian, 13 African American, 17 Hispanic, 25 multi-racial

appropriate EF model and highlights the importance of considering multiple model configurations to fully understand the structure of EF.

A key aspect of the studies is the number of factors considered. One-factor and two-factor models were most frequently tested. Specifically, twelve studies evaluated a one-factor model (Hume, 2015; Lerner & Lonigan, 2014; Lonigan et al., 2016; Miller et al., 2012; Miller et al., 2013; Monette et al., 2015; Mulder et al., 2014; Panesi & Morra, 2021; Usai et al., 2014; Völter et al., 2022; Wiebe et al., 2008; Wiebe et al., 2011), and an equal number investigated two-factor models (Hume, 2015; Lerner & Lonigan, 2014; Lonigan et al., 2016; Miller et al., 2012; Monette et al., 2015; Mulder et al., 2014; Miller et al., 2013; Panesi & Morra, 2021; Usai et al., 2014; Völter et al., 2022; Wiebe et al., 2008; Wiebe et al., 2011). Of these, six studies explored more than one two-factor model (Lerner & Lonigan, 2014; Miller et al., 2012; Miller et al., 2013; Monette et al., 2015; Völter et al., 2022; Wiebe et al., 2008). Additionally, the fit of three-factor models was analyzed in six studies (Lerner & Lonigan, 2014; Miller et al., 2012; Monette et al., 2015; Usai et al., 2014; Völter et al., 2022; Wiebe et al., 2008), with two of these studies examining alternative three-factor models (Völter et al., 2022; Wiebe et al., 2008). Remarkably, Völter and colleagues (2022) also tested a comprehensive nine-factor model, underscoring the complexity and breadth of their analytical approach.

It is noteworthy that among the studies supporting a unitary model, many only compared one-factor and two-factor models. Conversely, in the subset of research endorsing a two-factor model, four studies also evaluated three-factor models. This selective approach can potentially restrict the outcomes: CFA inherently depends on the range of models hypothesized by researchers. As such, the final model identified is intrinsically linked to the models included in the comparative analysis.

It is important to note that some studies investigating the validity of a three-factor model did not conceptualize three distinct EF (e.g., WM/updating, inhibition, and shifting); instead, they hypothesized that one function comprised multiple subcomponents for each EF. This approach was adopted for example by Wiebe and colleagues (2008) and Lerner and Lonigan (2014), both of whom considered various subcomponents of inhibition within their models. This suggests a nuanced understanding of EF, where single functions may have internal subdivisions that influence cognitive processing, but makes it hard to compare models from different articles because the number of tasks was often restricted making it hard to also test the presence of subcomponents within the three EFs.

2.3. Which latent variables received more attention?

The majority of studies predominantly focused on WM and inhibition. WM was featured in eleven studies (Hume et al., 2015; Lerner & Lonigan, 2014; Lonigan et al., 2016; Miller et al., 2012; Miller et al., 2013; Monette et al., 2015; Panesi & Morra, 2021; Usai et al., 2014; Völter et al., 2022; Wiebe et al., 2008; Wiebe et al., 2011), and inhibition was explored in ten (Hume et al., 2015; Lerner & Lonigan, 2014; Lonigan et al., 2016; Miller et al., 2012; Miller et al., 2013; Monette et al., 2015; Usai et al., 2014; Völter et al., 2022; Wiebe et al., 2008; Wiebe et al., 2011). Shifting, however, was considered in only four studies, always alongside WM and inhibition (Miller et al., 2012; Monette et al., 2015; Usai et al., 2014; Völter et al., 2022). One study examined a general EF factor independent of WM (Panesi & Morra, 2021), and another differentiated between 'cool' and 'hot' EFs (Mulder et al., 2014).

As previously discussed, the selection of latent variables significantly influences the results of the CFA. Notably, few studies incorporated shifting into their hypothesized models, which complicates direct comparisons across different studies. Despite the impact of the type and number of latent variables on the model outcomes, these factors alone are insufficient for definitive conclusions. Eight studies focused exclusively on WM and inhibition, but did not include shifting (Hume et al., 2015; Lerner & Lonigan, 2014; Lonigan et al., 2016; Miller et al., 2012; Miller et al., 2013; Wiebe et al., 2008; Wiebe et al., 2011). Intriguingly, these studies resulted in both one-factor and two-factor models.

2.4. How were the tasks employed actually informative of the constructs?

A significant issue that may account for the inconsistencies observed among studies concerning the latent structure of EF is the selection of tasks used for assessment. The type, number, and association of indicators can produce different results.

First of all, we can notice an extremely high variability in the tasks employed to measure the same latent variables. Most measures were employed only in one study, with few exceptions: for example, the Backward Word Span (Miller et al., 2012; Miller et al., 2013; Monette et al., 2015) and the Stroop Day/Night Task (Hume, 2015; Lerner & Lonigan, 2014; Lonigan et al., 2016; Monette et al., 2015) were the tasks which were employed most often and consistently across different studies.

Analyzing the tasks used across studies is complicated by the frequent use of different names for very similar tasks. For instance, the task named "Word Span Reversed" by Lerner and Lonigan (2014) and Lonigan et al. (2016) is essentially the same as "Backward Word Span," despite the naming difference. Similarly, Panesi and Morra (2021) refer to an inhibitory task as "Circle Tracing," whereas Usai and colleagues (2014) call a comparable task "Circle Drawing Task." This pattern extends to Stroop tasks like the Day-Night, Knock Tap, and other variants, which all involve inhibiting a prepotent response in favor of an incongruent one but use different stimuli to elicit the Stroop effect. Furthermore, tasks that are based on the span process, such as Word Span Reverse, Listening Span, and others, require recalling an increasing number of elements. These tasks often employ different materials, adding another layer of complexity. Additionally, the format of the tests varies, with some conducted on paper and others computerized, which can influence the task outcomes. Moreover, even when the same task is used, most studies adopt different scoring procedures, making direct comparisons between tasks used in different studies even more challenging.

A significant methodological issue arises when the same task is employed to assess different EF factors across studies. For instance, the Preschool Continuous Performance Test (CPT) was used to evaluate both WM and inhibition; Miller et al. (2013) used it to measure WM, while Miller et al. (2012) and Wiebe et al. (2008) used it for assessing inhibition. Notably, Miller et al. (2012) used the Preschool

CPT to indicate both WM and inhibition within the same study. Similarly, the Go-No Go Task was utilized by the same group to assess both shifting and inhibition. This issue extends to different EF factors being measured by the very same tasks. For example, Panesi and Morra (2021) used the Circle Tracing task to assess a generic EF factor, whereas Usai et al. (2014) used it specifically for inhibition. The Six Boxes task was another example, employed both to measure cool EF in the study by Mulder et al. (2014) and WM by Wiebe et al. (2008). Furthermore, the Snack Delay task was used to assess both hot EF and inhibition in different studies. Such practices can complicate the interpretation of results, as they may not only reflect the intended cognitive processes but also overlapping or unrelated cognitive demands.

Furthermore, some of the studies included in the present systematic review used tasks measuring abilities often different from those that they are traditionally intended to. For example, Miller and colleagues (2012; 2013) and Wiebe and colleagues (2008) employed the Tower of Hanoi, and Usai and colleagues used the Tower of London. These two tasks are generally used as a measure of planning, while in these cases they were employed as measures of inhibition. This may be due to the fact that young children's performance in these tasks is significantly affected by children's ability to respect the rules, rather than their emerging planning ability (see also Bull, Espy, & Senn, 2004). This was suggested by other empirical evidence, like Lehto and colleagues (2003) and Senn et al. (2004). This evidence suggests that children may approach the task in different ways, making it hard to understand how the use of certain tasks is informative of the constructs they would assess.

Table 2 illustrates that there are considerable differences in the number of indicators used to measure each latent variable in the studies reviewed. A few studies used only two indicators, the minimum required for inclusion in this review (Monette et al., 2015; Mulder et al., 2014; Usai et al., 2014). Most studies employed three or four indicators (Hume, 2015; Lerner & Lonigan, 2014; Lonigan et al., 2016; Miller et al., 2012; Miller et al., 2013; Monette et al., 2015; Mulder et al., 2014; Panesi & Morra, 2021; Völter et al., 2022; Wiebe et al., 2008; Wiebe et al., 2011). Notably, two studies used more than four indicators to measure EF, Lonigan and colleagues (2016) used six indicators and Wiebe et al. (2008) used seven indicators to assess inhibition. It is also worth noting that no study used the same number of indicators uniformly for each EF component. For instance, eight studies varied in the number of indicators for each EF component (Lerner & Lonigan, 2014; Lonigan et al., 2016; Miller et al., 2012; Monette et al., 2015; Mulder et al., 2014; Panesi & Morra, 2021; Wiebe et al., 2008; Wiebe et al., 2011), with Wiebe et al. (2008) presenting the widest range—three indicators for WM and as many as seven for measuring inhibition. Only four studies included in this review included a consistent number of indicators for each EF factor (Hume, 2015; Miller et al., 2013; Usai et al., 2014; Völter et al., 2022). This wide variability in the choice and number of indicators highlights the challenges in comparing results across different studies and emphasizes the need for more consistent methods for assessing EF.

2.5. How does the range of age affect the structure of EF?

Table 2 clearly shows that the age ranges of preschooler samples in the studies are not uniform. Four studies encompassed the entire preschool range of 3–6 years (Lonigan et al., 2016; Miller et al., 2012; Miller et al., 2013; Völter et al., 2022), while Wiebe et al. (2008) extended their study to include children from 2 to 6 years. Other studies targeted narrower age spans; for instance, Usai et al. (2014) focused on children between 5 and 6 years, Lerner & Lonigan (2014) on children between 4 and 5 years, and two studies (Mulder et al., 2014; Panesi & Morra, 2021) included children between 2 and 3 years. Hume (2015) examined two distinct groups, one aged 3–4 years and another aged 4–5 years. Notably, Monette et al. (2015) did not specify the age range of their sample, whereas Wiebe et al. (2011) reported that children were assessed within three weeks of their third birthday.

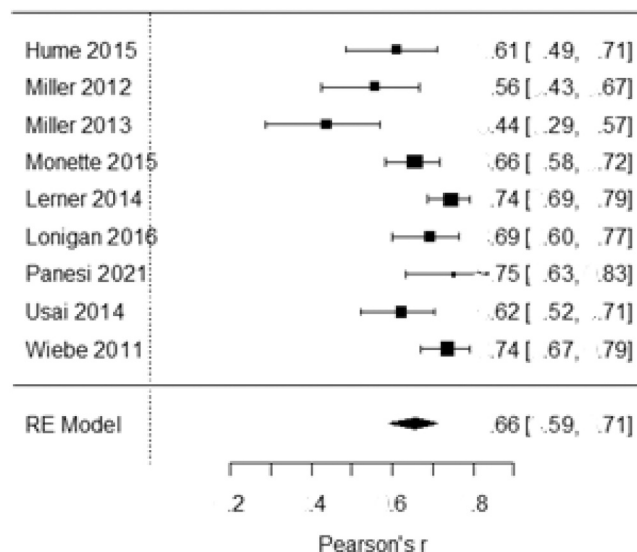


Fig. 2. Forest plot. The correlation on the bottom row indicates the overall meta-analytic effect.

It is evident from the review that all studies resulting in a one-factor model involved children up to 4 years. Conversely, studies that identified a two-factor model generally included children aged four to five years. The only exception was the study by Mulder and colleagues (2014): in this particular paper the authors found a distinction between diverse EF in younger children, but this distinction was between cool and hot EF. In sum, among studies which used the same classification of EF it seems clear that EF are a unique factor in children aged 3, and separated in two factors, which distinguish WM and inhibition, in 5-year-old children; four years may be considered as a transition age.

2.6. Meta-analysis

The meta-analytic estimate was $r = .66$, 95 % CI [.59,.71] indicating that there was a strong positive relationship between WM and inhibition. Fig. 2 shows the correlation between inhibition and WM in the different studies. On the bottom row of the figure is reported the overall meta-analytic effect. There was significant variability among the effects, $Q(8) = 33.67$, $p < .001$. The funnel plot, which is a scatter plot of individual studies showing their precision and results (each dot representing a single study), is provided in Fig. 3.

We also investigated the potential influence of publication bias, by conducting both a trim-and-fill analysis and Egger's regression test for funnel plot asymmetry. The trim-and-fill analysis estimated that no studies were missing on the right side of the funnel plot (estimated number of missing studies = 0, SE = 1.96), and Egger's test did not reveal significant asymmetry ($z = -0.68$, $p = .497$). These results suggest that the likelihood of publication bias affecting the meta-analytic estimate is low. Additionally, while the test for heterogeneity was statistically significant ($Q(8) = 33.67$, $p < .0001$), the estimated between-study variance was relatively small ($\tau^2 = 0.0203$, $\tau = 0.14$), indicating that the variation in effect sizes across studies was modest. Taken together, these findings support the robustness of our results and suggest that they are not substantially affected by publication bias or outlier effects.

3. General discussion

The structure of executive functions (EF) in adulthood is widely supported by the three-component model proposed by Miyake and colleagues, which has become well-established (Miyake et al., 2000; Miyake & Friedman, 2012). However, consensus on the EF structure during childhood remains elusive. Developmental studies yield conflicting results, particularly for preschool-aged children: some researchers advocate for a unitary structure of EF, while others support models with two distinct components (see Karr et al., 2018). The main goal of this systematic review is to critically examine the structure of EF in preschool children. Our analysis focuses on the factorial models analyzed, the EF components considered, the number and types of tasks used for assessment, and the age range of the study samples. Additionally, this review attempts to meta-analyze the latent correlations between WM and inhibition, aiming to clarify these relationships in early childhood.

As expected, our review uncovered several inconsistencies in the existing literature. Nearly half of the studies supported a one-factor model, while the rest advocated for a two-factor model. This division underscores the difficulty in identifying a universally accepted model, prompting a reflection on the caveats that significantly impact the outcomes of various studies. We noted substantial

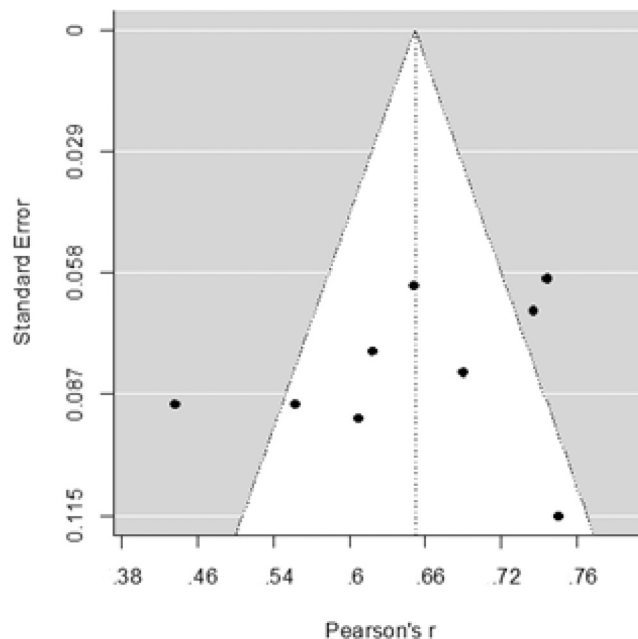


Fig. 3. Funnel plot.

variations across studies, particularly in the type and number of models tested; there was a prevalent comparison between one- or two-factor models among studies favoring a single-factor model, and an introduction of three-factor models in those identifying a two-factor model. Regarding the choice of latent variables, there was more consensus: most studies included WM and inhibition, while a subset adhered to Miyake's classical model from 2000, also incorporating shifting. However, a minority of studies explored different operationalizations of EF, such as the distinction between cool EF and hot EF, following the framework proposed by Zelazo and Muller (2002).

A critical issue we observed is the high variability in the tasks employed to measure the same latent factors across different studies, often using very distinct tasks. This variability, coupled with ambiguities such as similar tests presented under different names, can significantly affect the results. Furthermore, the use of tests far removed from their original intended purposes adds to this complexity (e.g., using a task originally intended to measure planning to measure inhibition). Another problem relies to the use of very similar tasks to measure the same factor. For example, Miller and colleagues (2013) assessed WM using both Backward Word Span and Backward Digit Span. The high correlation observed between these tasks—due to their shared nature as span tasks involving verbal inputs—may not truly reflect the underlying latent variable variance (Cole, Ciesla, & Steiger, 2007). This suggests that employing a broader range of measures could mitigate such issues. Additionally, the problem of task impurity arises when tasks designed to measure specific cognitive processes require multiple cognitive functions. This impurity can lead to skewed conclusions about the specific cognitive processes being examined. Controlling for task impurity is not easy, because any complex task unavoidably involves multiple cognitive abilities and functions. This issue becomes particularly relevant when studying preschool children, as their task performance is influenced not only by executive processes but also by other developmental factors such as language skills, attention span, motivation, and familiarity with structured testing situations. Moreover, because of the more limited behavioral repertoire of preschool children, the tasks used in this age group are necessarily simpler, and may therefore be more homogeneous in their executive demands compared to those used with older children or adults (Wiebe et al., 2008). Careful task analysis guided by sound cognitive theory could sometimes help identify which tasks share the involvement of not only a specific EF but also other knowledge or abilities. Thus, researchers could posit in their CFA models correlations between the error variances of specific task pairs, or even posit additional factors that represent content variance shared by a number of tasks. In this way, the problem of task impurity could be mitigated, although fully controlling it seems very difficult. Nevertheless, acknowledging these limitations is essential to appropriately interpret the structure of EF in early childhood, while also recognizing the need to combine insightful, theoretically informed modelling with methodological caution.

The issue of which measures should be used as indicators of each latent variable is exacerbated in studies like those by Miller et al. (2012), Miller et al. (2013), and Monette et al. (2015), which utilize the same tasks but different scoring methods in studies. The highly correlated errors from these indices can complicate result interpretation (Kline, 2015). Additionally, problems arise when tasks are allowed to load on different latent factors, as seen in Monette et al. (2015) and Miller et al. (2012), where indicators from the same tasks are loaded onto different latent variables. This method can obscure the distinction between EF factors.

Given the aforementioned issues, it is important to consider whether CFA is always the appropriate analysis for identifying the latent structure of EF, if the measures included are too strongly correlated or pertaining to the very same task, for example. CFA starts with a theoretical model that highlights the relationships among the variables in a given model, commonly referred to as the measurement model. This model is then compared against the observed data to evaluate the model's fit; this involves assessing how well the covariance matrix, as estimated from the available data, replicates the observed covariance matrix. However, if there are methodological biases related to the choice of indicators or variables, the results from CFA might not be entirely reliable. These biases could significantly influence the model, and by extension, the interpretations drawn about the latent structures of EF. This question is highlighted by Miller et al. (2012): in their study, they tested two different sets of models: firstly, they tried to replicate the structure found by Wiebe et al. in 2008, administering WM and inhibition tasks; in this case, they accepted a one-factor model. Secondly, they added some performance indicators to distinctly measure WM, shifting, and inhibition factors. As result, a two-factor model consisting of a WM-shifting factor and an inhibition one fit the data better than a single-factor model. Thanks to these findings, the authors suggested that the structure of EF in preschoolers that emerges from CFA was influenced by task and performance indicator selection.

Concerning the age range, we observed that the age range was very heterogeneous across different studies. Moreover, while some studies focused on a specific age range (Wiebe et al., 2011) other considered a larger one (Wiebe et al., 2008). Our review indicates that studies with samples of children under 4 years typically identified a one-factor model, whereas those with children older than 4 years found two-factor models. These results imply that EFs may differentiate as children develop, with the age around four and a half years potentially marking a critical transition stage from a unitary to a bifurcated EF structure (see Panesi and Morra, 2020 for a possible theoretical explanation).

The results of our meta-analysis provide additional insights into the structure of EF in preschool children, particularly through an analysis of the correlations between inhibition and WM at the latent level. A significant correlation of .66 was observed, underscoring a strong relationship between these two EF components. This finding corroborates the observations by Panesi and colleagues (2020), who noted a particularly strong link between WM and inhibition. This correlation might be explained by the theoretical model proposed by Im-Bolter, Johnson, and Pascual-Leone (2006), which posits that M capacity and inhibition are overarching resources that develop concurrently and interactively throughout childhood. Additionally, this result supports Miyake and colleagues' (2000) hypothesis that all EF components are, to some extent, dependent on WM. Our findings emphasize that while WM and inhibition are distinct constructs, they are also highly interrelated. From the perspective of latent structure, this relationship is particularly meaningful: understanding EF in preschool children involves not only identifying which latent components are present, but also examining the strength of the associations between them. The observed correlation of .66 between WM and inhibition indicates a strong relationship, suggesting that these functions are closely connected in this age range. However, this value is not so high as to imply

redundancy or overlap between the constructs, thereby supporting the idea that they represent distinguishable but interconnected components of executive functioning. Understanding this relationship is crucial for identifying which cognitive abilities are already developed in preschool children and how they interact.

This finding is particularly informative because it helps clarify what can be expected from children of this age regarding these functions. Specifically, we can anticipate that children with stronger WM abilities will also demonstrate better inhibitory control, while those with weaker WM may struggle more with inhibition. This is in line with the findings of recent studies, including Lerner and Lonigan (2014). Recognizing this link provides valuable insights into typical EF development, which can be instrumental in shaping both assessment strategies and intervention programs aimed at supporting cognitive growth in early childhood. Furthermore, these results highlight the need for future research to account for the strong interdependence between WM and inhibition when examining EF structure. Investigating these components in isolation may overlook their dynamic interaction, which plays a crucial role in shaping cognitive development. Additionally, a deeper exploration of this relationship in different populations, including children with neurodevelopmental conditions or those from non-WEIRD countries, could provide further insights into both typical and atypical EF trajectories. Understanding how WM and inhibition interact in diverse developmental contexts may refine theoretical models and improve targeted interventions for children with EF difficulties.

While this study has yielded important insights, it also highlights several areas for improvement in future research. To broaden our understanding of EF in preschoolers, it is crucial to expand beyond primarily measuring inhibition and WM. Including tasks that assess shifting—a component seldom explored in this age group—could provide a more comprehensive view of EF. The choice of tasks used in CFA is critical. Future studies should avoid using tasks that are intrinsically highly correlated or employing different scores from the same task within their models, as this can complicate the interpretation of the CFA. Although using two indicators per latent factor is common, this can lead to problems in the estimation of the latent variables for these factors (Kline, 2015). A more extensive array of tasks, ideally with a consistent number of indicators for each latent factor, would enhance model reliability and validity. Moreover, employing more than two indicators for each latent variable and avoiding similar tasks could improve the robustness of the findings.

Careful consideration of the age range of participants is also essential; employing a large, comprehensive battery of tests across a broad cohort of children for each individual age can significantly aid in understanding the nuanced structure of EF in preschoolers. Finally, as developmental changes are continually occurring, conducting longitudinal studies would be invaluable. These studies could identify when and how relationships between factors change, and pinpoint critical periods in the development of EF.

It is also important to highlight some limits of the current review. Regrettably our review was limited by the paucity of research in this area yielding to a small number of studies considered. Our inclusion criteria were intentionally stringent to ensure a reasonably homogeneous set of studies. We specifically focused on research that examines the latent structure of EF using CFA, widely regarded as the most effective method for confirming the structure of EF (Miyake et al., 2000). However, different approaches, such as Exploratory Factor Analysis (EFA), Bifactor Modeling, Latent Profile Analysis (LPA), or Multilevel Modeling, may have allowed to consider other studies on EF structure from both white and grey literature. In this regard, it is important to acknowledge that alternative conceptualizations of EF exist, along with different analytic strategies for exploring their nature (e.g., Camerota, Willoughby, & Blair, 2020; Doebel, 2020; Willoughby, Holochwost, Blanton, & Blair, 2014). Future reviews should consider including studies that address these broader perspectives to provide a more comprehensive understanding of EF across various age groups and theoretical frameworks. The meta-analysis conducted was based on a limited number of studies, which is not ideal. We recommend that future meta-analyses include a larger pool of studies to enhance the robustness of the findings. However, the availability of a more extensive dataset is contingent upon the publication of additional research focused on preschool children, which is currently sparse. It is also important to note that the small sample size in our analysis precluded the testing for potential moderators (e.g., age).

In conclusion, this systematic review has highlighted the complex and multifaceted nature of EF in preschool children. Our findings reveal significant variability in the conceptualization and measurement of EF, influenced heavily by the choice of tasks, the models employed, and the age of the participants. Therefore, although some studies suggest a unitary structure of EF among younger preschoolers and others point towards a differentiated model as children approach school age, future studies are needed to verify whether these results are confirmed even when different constructs are included (e.g. inhibition, shifting, updating, WM) and diverse methodological approaches are employed. Additionally, our meta-analysis of the latent correlations between WM and inhibition contributes to a deeper understanding of these functions' interdependencies. As we move forward, it is crucial that future studies refine EF measurement strategies, expand the range of considered variables, and embrace longitudinal designs to track changes across time.

These findings are not only of theoretical interest but also have important practical implications. A better understanding of the organization of EF provides insight into which abilities are already developed in early childhood and which are still emerging, clarifying which abilities children use to complete specific tasks. This is especially important in the preschool years, a period of rapid cognitive change. Furthermore, a more precise understanding of the latent structure of EF enables a deeper investigation of its interaction with socio-cognitive abilities, such as learning, emotional competence, and theory of mind. From an educational perspective, this understanding can inform classroom-based programs and enhance teachers' awareness of which skills to foster and how they influence other developmental domains. Clinically, a clearer picture of typical EF development is essential for distinguishing between typical and atypical trajectories. These insights also have significant implications for EF assessment, helping to determine which abilities should be evaluated and which tests should be used to ensure accurate measurement. Finally, they are crucial for designing effective intervention strategies: knowing the structure of EF allows for targeted stimulation of specific skills, supporting cognitive and socio-emotional development in young children.

CRedit authorship contribution statement

David Giofrè: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation. **Sabrina Panesi:** Writing – review & editing, Writing – original draft, Supervision. **Sergio Morra:** Writing – review & editing, Writing – original draft, Supervision. **Laura Traverso:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Data curation, Conceptualization. **Alessia Bandettini:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Data availability

Data will be made available on request.

References

- Baddeley, A. (1986). *Working memory*. Oxford University Press.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829. <https://doi.org/10.1038/nrn1201>
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>
- Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, 78(2), 647–663. <https://doi.org/10.1111/j.1467-8624.2007.01019.x>
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. John Wiley & Sons, Ltd.
- Camerota, M., Willoughby, M. T., & Blair, C. B. (2020). Measurement models for studying child executive functioning: questioning the status quo. *Developmental Psychology*, 56(12), 2236. <https://doi.org/10.1037/dev0001127>
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, 28(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3
- Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, 12(4), 381–398. <https://doi.org/10.1037/1082-989X.12.4.381>
- Conklin, H. M., Salorio, C. F., & Slomine, B. S. (2008). Working memory performance following paediatric traumatic brain injury. *Brain Injury*, 22(11), 847–857. <https://doi.org/10.1080/02699050802403565>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science*, 15(4), 942–956. <https://doi.org/10.1177/1745691620904771>
- Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86, 186–204. <https://doi.org/10.1016/j.cortex.2016.04.023>
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, 134(1), 31–60. <https://doi.org/10.1037/0033-2909.134.1.31>
- Garon, N., Smith, I. M., & Bryson, S. E. (2014). A novel executive function battery for preschoolers: Sensitivity to age differences. *Child Neuropsychology*, 20(6), 713–736. <https://doi.org/10.1080/09297049.2013.857650>
- Godfrey, O., Cabaret, M., Petit-Chenal, V., Pruvo, J. P., & Rousseaux, M. (1999). Control functions of the frontal lobes: Modularity of the central-supervisory system? *Cortex*, 35(1), 1–20. [https://doi.org/10.1016/s0010-9452\(08\)70782-2](https://doi.org/10.1016/s0010-9452(08)70782-2)
- Himi, S. A., Bühner, M., & Hilbert, S. (2021). Advancing the understanding of the factor structure of executive functioning. *Journal of Intelligence*, 9(1), 16. <https://doi.org/10.3390/jintelligence9010016>
- Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: predictive relations from ages 2 to 4 years. *Developmental Psychology*, 43(6), 1447–1459. <https://doi.org/10.1037/0012-1649.43.6.1447>
- Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, 35(1), 20–36. <https://doi.org/10.1080/87565640903325691>
- Hughes, C., & Graham, A. (2002). Measuring executive functions in childhood: Problems and solutions? *Child and Adolescent Mental Health*, 7(3), 131–142. <https://doi.org/10.1111/1475-3588.00024>
- Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variables analysis. *Neuropsychologia*, 44, 2017–2036. <https://doi.org/10.1016/j.neuropsychologia.2006.01.010>
- Hume, L. E. (2015). *The structure and nature of executive functions in 3- and 4-year-old children* [Doctoral dissertation, Florida State University]. (http://purl.flvc.org/fsu/fd/FSU_migr_etd-9617)
- Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development*, 77, 1822–1841. <https://doi.org/10.1111/j.1467-8624.2006.00976.x>
- Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, 144(11), 1147–1185. <https://doi.org/10.1037/bul0000160>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Lee, K., Bull, R., & Ho, R. M. H. (2013). Developmental changes in executive functioning. *Child Development*, 84, 1933–1953. <https://doi.org/10.1111/cdev.12096>
- Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21, 59–80. <https://doi.org/10.1348/026151003321164627>
- Lerner, M. D., & Lonigan, C. J. (2014). Executive function among preschool children: Unitary versus distinct abilities. *Journal of Psychopathology and Behavioral Assessment*, 36, 626–639. <https://doi.org/10.1007/s10862-014-9424-3>
- Lonigan, C. J., Lerner, M. D., Goodrich, J. M., Farrington, A. L., & Allan, D. M. (2016). Executive function of Spanish-speaking language-minority preschoolers: structure and relations with early literacy skills and behavioral outcomes. *Journal of Experimental Child Psychology*, 144, 46–65. <https://doi.org/10.1016/j.jecp.2015.11.003>
- McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, 16(3), 495–505. <https://doi.org/10.1017/S1355617710000093>
- Miller, M. R., Giesbrecht, G. F., Muller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, 13, 395–423. <https://doi.org/10.1080/15248372.2011.585478>
- Miller, M. R., Müller, U., Giesbrecht, G. F., Carpendale, J. I., & Kerns, K. A. (2013). The contribution of executive function and social understanding to preschoolers' letter and math skills. *Cognitive Development*, 28(4), 331–349. <https://doi.org/10.1016/j.cogdev.2012.10.005>
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1), 8–14. <https://doi.org/10.1177/0963721411429458>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>

- Monette, S., Bigras, M., & Lafrenière, M. A. (2015). Structure of executive functions in typically developing kindergartners. *Journal of Experimental Child Psychology*, 140, 120–139. <https://doi.org/10.1016/j.jecp.2015.07.005>
- Morra, S., Panesi, S., Traverso, L., & Usai, M. C. (2018). Which tasks measure what? Reflections on executive function development and a commentary on Podjarny, Kamawar, and Andrews 2017. *Journal of Experimental Child Psychology*, 167, 246–258. <https://doi.org/10.1016/j.jecp.2017.11.004>
- Morris, N., & Jones, D. M. (1990). Memory updating in working memory: the role of the central executive. *British Journal of Psychology*, 81, 111–121. <https://doi.org/10.1111/j.2044-8295.1990.tb02349.x>
- Mulder, H., Hoofs, H., Verhagen, J., van der Veen, I., & Leseman, P. P. (2014). Psychometric properties and convergent and predictive validity of an executive function test battery for two-year-olds. *Frontiers in Psychology*, 5, 733. <https://doi.org/10.3389/fpsyg.2014.00733>
- Norman, D. A., & Shallice, T. (1980). Attention to action: Willed and automatic control of behavior (Technical Report No. 99). *Center for Human Information Processing, University of California, Sanoma Diego*.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory*, 4 pp. 1–18). US: Springer.
- Panesi, S., Bandettini, A., Traverso, L., & Morra, S. (2022). On the relation between the development of working memory updating and working memory capacity in preschoolers. *Journal of Intelligence*, 10(1), 5. <https://doi.org/10.3390/jintelligence10010005>
- Panesi, S., & Morra, S. (2020). Executive functions and mental attentional capacity in preschoolers. *Journal of Cognition and Development*, 21(1), 72–91. <https://doi.org/10.1080/15248372.2019.1685525>
- Panesi, S., & Morra, S. (2021). Executive function, language, and the toddler's discovery of representational drawing. *Frontiers in Psychology*, 12, Article 659569. <https://doi.org/10.3389/fpsyg.2021.659569>
- Schoemaker, K., Bunte, T., Wiebe, S. A., Espy, K. A., Deković, M., & Matthys, W. (2012). Executive function deficits in preschool children with ADHD and DBD. *Journal of Child Psychology and Psychiatry*, 53(1), 111–119. <https://doi.org/10.1111/j.1469-7610.2011.02468.x>
- Senn, T. E., Espy, K. A., & Kaufmann, P. M. (2004). Using path analysis to understand executive function organization in preschool children. *Developmental Neuropsychology*, 26(1), 445–464. https://doi.org/10.1207/s15326942dn2601_5
- Shing, Y. L., Lindenberger, U., Diamond, A., Li, S. C., & Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Developmental Neuropsychology*, 35(6), 679–697. <https://doi.org/10.1080/87565641.2010.508546>
- Sølsnes, A. E., Skranes, J., Brubakk, A. M., & Løhaugen, G. C. (2014). Executive functions in very-low-birth-weight young adults: a comparison between self-report and neuropsychological test results. *Journal of the International Neuropsychological Society*, 20(5), 506–515. <https://doi.org/10.1017/S1355617714000332>
- Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin*, 95(1), 3–28. <https://doi.org/10.1037/0033-2909.95.1.3>
- Tranel, D., Anderson, S. W., & Benton, A. (1994). Development of the concept of “executive function” and its relationship to the frontal lobes. In In. F. Boller, & J. Grafman (Eds.), *Handbook of neuropsychology*, 9 pp. 125–148). Elsevier.
- Usai, M. C., Viterbori, P., Traverso, L., & De Franchis, V. (2014). Latent structure of executive function in five- and six-year-old children: a longitudinal study. *European Journal of Developmental Psychology*, 11, 447–462. <https://doi.org/10.1080/17405629.2013.840578>
- van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, Reading, and arithmetic. *Intelligence*, 35, 427–449. <https://doi.org/10.1016/j.intell.2006.09.001>
- Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, 36(3). <https://doi.org/10.18637/jss.v036.i03>
- Völter, C. J., Reindl, E., Felsche, E., Civelek, Z., Whalen, A., Lugosi, Z., Duncan, L., Herrmann, E., Call, J., & Seed, A. M. (2022). The structure of executive functions in preschool children and chimpanzees. *Scientific Reports*, 12(1), 6456. <https://doi.org/10.1038/s41598-022-08406-7>
- Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, 44(2), 575–587. <https://doi.org/10.1037/0012-1649.44.2.575>
- Wiebe, S., Sheffield, T., Nelson, J. M., Clark, C. A. C., Chevalier, N., & Espy, K. (2011). The structure of executive function in 3-year-olds. *Journal of Experimental Child Psychology*, 108, 436–452. <https://doi.org/10.1016/j.jecp.2010.08.008>
- Willoughby, M., Holochwost, S. J., Blanton, Z. E., & Blair, C. B. (2014). Executive functions: Formative versus reflective measurement. *Measurement: Interdisciplinary Research Perspectives*, 12(3), 69–95. <https://doi.org/10.1080/15366367.2014.929453>
- Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology*, 1(2), 198–226. <https://doi.org/10.1037/1089-2680.1.2.198>
- Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In In. U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 445–470). Blackwell Publishing.